

1/6

Input file flh14926cons; Output File flh14926tra  
Sequence length 2818

```
CCACGCGTCCGATTATAGAAAGAAGCTATAGTCTGTCTGGGATATCCAACACCAACAGGGCTTACTGAGAGCTCCATTT
CTGGAAAGCCTTACAAGACTGAGGAATATCAGACTGCGAATCACCAGGAACGGTTCCTTGACGACAGAAGCAATCTC
TCTCCCATCTTCGCATATTCTAATGGCAAAACAAGTGAAGAAAAGAGGAAGCATGACTGCAGATCAGATCAGTTCTC
TTTGTGGATTATATTTTCAGTAAATGTATGGATCTATCTTTTCTTGTTCTTATATCTAGATCATGAGACTTGACTGA

      M A N Y S H A A D N I L Q      13
GGCTGTATCCTTATCCTCCATCCATCT ATG GCG AAC TAT AGC CAT GCA GCT GAC AAC ATT TTG CAA      39
N L S P L T A F L K L T S L G F I I G V      33
AAT CTC TCG CCT CTA ACA GCC TTT CTG AAA CTG ACT TCC TTG GGT TTC ATA ATA GGA GTC      99
S V V G N L L I S I L L V K D K T L H R      53
AGC GTG GTG GGC AAC CTC CTG ATC TCC ATT TTG CTA GTG AAA GAT AAG ACC TTG CAT AGA      159
A P Y Y F L L D L C C S D I L R S A I C      73
GCA CCT TAC TAC TTC CTG TTG GAT CTT TGC TGT TCA GAT ATC CTC AGA TCT GCA TTT TGT      219
F P F V F N S V K N G S T W T Y G T L T      93
TTC CCA TTT GTG TTC AAC TCT GTC AAA AAT GGC TCT ACC TGG ACT TAT GGG ACT CTG ACT      279
C K V I A F L G V L S C F H T A F M L F      113
TGC AAA GTG ATT GCC TTT CTG GGG GTT TTG TCC TGT TTC CAC ACT GCT TTC ATG CTC TTC      339
C I S V T R Y L A I A H H R F Y T K R L      133
TGC ATC AGT GTC ACC AGA TAC TTA GCT ATC GCC CAT CAC CGC TTC TAT ACA AAG AGG CTG      399
T F W T C L A V I C M V W T L S V A M A      153
ACC TTT TGG ACG TGT CTG GCT GTG ATC TGT ATG GTG TGG ACT CTG TCT GTG GCC ATG GCA      459
F P P V L D V G T Y S F I R E E D Q C T      173
TTT CCC CCG GTT TTA GAC GTG GGC ACT TAC TCA TTC ATT AGG GAG GAA GAT CAA TGC ACC      519
F Q H R S F R A N D S L G F M L L L A L      193
TTC CAA CAC CGC TCC TTC AGG GCT AAT GAT TCC TTA GGA TTT ATG CTG CTT CTT GCT CTC      579
I L L A T Q L V Y L K L I F F V H D R R      213
ATC CTC CTA GCC ACA CAG CTT GTC TAC CTC AAG CTG ATA TTT TTC GTC CAC GAT CGA AGA      639
K M K P V Q F V A A V S Q N W T F H G P      233
AAA ATG AAG CCA GTC CAG TTT GTA GCA GCA GTC AGC CAG AAC TGG ACT TTT CAT GGT CCT      699
G A S G Q A A A N W L A G F G R G P T P      253
GGA GCC AGT GGC CAG GCA GCT GCC AAT TGG CTA GCA GGA TTT GGA AGG GGT CCC ACA CCA      759
P T L L G I R Q N A N T T G R R R L L V      273
CCC ACC TTG CTG GGC ATC AGG CAA AAT GCA AAC ACC ACA GGC AGA AGA AGG CTA TTG GTC      819
L D E F K M E K R I S R M F Y I M T F L      293
TTA GAC GAG TTC AAA ATG GAG AAA AGA ATC AGC AGA ATG TTC TAT ATA ATG ACT TTT CTG      879
F L T L W G P Y L V A C Y W R V F A R G      313
TTT CTA ACC TTG TGG GGC CCC TAC CTG GTG GCC TGT TAT TGG AGA GTT TTT GCA AGA GGG      939
```

TO FIG. 1B.

FIG. 1A.

2/6

FROM FIG. 1A.

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |      |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|------|
| P   | V   | V   | P   | G   | G   | F   | L   | T   | A   | A   | V   | W   | M   | S   | F   | A   | Q   | A   | G   |  | 333  |
| CCT | GTA | GTA | CCA | GGG | GGA | TTT | CTA | ACA | GCT | GCT | GTC | TGG | ATG | AGT | TTT | GCC | CAA | GCA | GGA |  | 999  |
| I   | N   | P   | F   | V   | C   | I   | F   | S   | N   | R   | E   | L   | R   | R   | C   | F   | S   | T   | T   |  | 353  |
| TTT | CCC | CCG | GTT | TTA | GAC | GTG | GGC | ACT | TAC | TCA | TTC | ATT | AGG | GAG | GAA | GAT | CAA | TGC | ACC |  | 1059 |
| L   | L   | Y   | C   | R   | K   | S   | R   | L   | P   | R   | E   | P   | Y   | C   | V   | I   | *   |     |     |  | 371  |
| CTT | CTT | TAC | TGC | AGA | AAA | TCC | AGG | TTA | CCA | AGG | GAA | CCT | TAC | TGT | GTT | ATA | TGA |     |     |  | 1113 |

GGGAGCATCTGTAAATCTTTAGCCTTGTGAAACTAACCTTCTCTGCTGAGCAATTGTGGCCATAGCCATATTTTGAG  
 AAGGAAATCAAGAATGGAATCAGCAGTTTAAAGGATTTGGGCAACATTCTGCAGTCTTTGCAATAGTTACCTATAATC  
 CTATTTTAAATCTCAGAGTGATCCTGCTGACTGCCAGCAAAGGTTTGTAAATTAAGAAGGGACTGAACCACTGCCCTAAG  
 TTTCTTTTGTGGTCAAAAACCTAGATAATGAAAGTAGCAGGTGCTAAGTATCAGTGCTAAATGCTCTGTATGTCACTAC  
 ATATGAAAAACATAAAAAACAATTAGCATTGGACATCTTAATAAATTAAGTTGACATGAGGTAAATGTGTTGATAAA  
 AACTAATTTTGAAGTTTGAAGACTTTAAACATTTCTACTACTATTGTTTTGCAAAGACTAAAATATTTGGGGACTT  
 AAAGTACTGTAATCCACTAAAGACGTGCCAATGAATTATTGGAATATCACACTTTAAAAACCGCCTTGTAAGTTCTGGG  
 GAGCATTCCAAAGCAGTATATTGGTTCCAATTAGAGTTTACTTTTTTTGTATTAATACATTGCTATTTCTAAATACCAC  
 TTTCTCATCTACTAGTAAGATTGCTAGCATTGAACTGTATTATGTGGTTTTTTGTTGATTGGTATAAAGTTTTTCCAA  
 TTCATTTATATTTTACAATGCTAGATATTGGTCTGGGAGGCAACATTAATGGTACCAGCCTGTCACAACTGAGCAGTT  
 CTAATAATGCAGAATAAACACATGTTGCCTTAAAGGGTTATCTAGKATCCYTTCTCTTATTAGCACTGGAGCAAATAG  
 YCAAGGGAAATCRAATCAGTAAGTGGTCATGGTCATGCATCTRAAAGTGCATGGAAGATCATTAGTACTTTTTCCCTT  
 TTTCTCACATGGTTTGAACTTAAAGTGCACATCMCTGAAATAATGAGATTTTCTTTTRMGGTGTGCTACCTTYYTAR  
 ASTGTTCTAAGAAGCAGGAGTTGATGTATGTTTATATTTTAAAGTCAGCTGTCGAGGGGAGACCACAGCCCTTAGTATGA  
 CATCCTGCACAATTTGTGAAGCATTATTTCTACTGAAGGCACAGTCTTGTTTATACTTTCTGCACATTCAGTGTATTGG  
 TCATTTAAATTATTTAGTTTTAACTTGTAAGCTTATAATATGATTTCTGGTATTTTAGAAATACATTAGAGTCTGT  
 GAGTCTCATTCTTTAAGATACANATGTGTGAACCTCAATATAAAGTTGCATTTGCCAAAATTTAAAAAAAAAAAAAAAAA  
 AAAAAAAAAA

FIG. 1B.

3/6

**Protein Family/Domain HMM Matches for f1h14926orfaa**

>PF00001/7th\_1 7 transmembrane receptor (rhodopsin family)

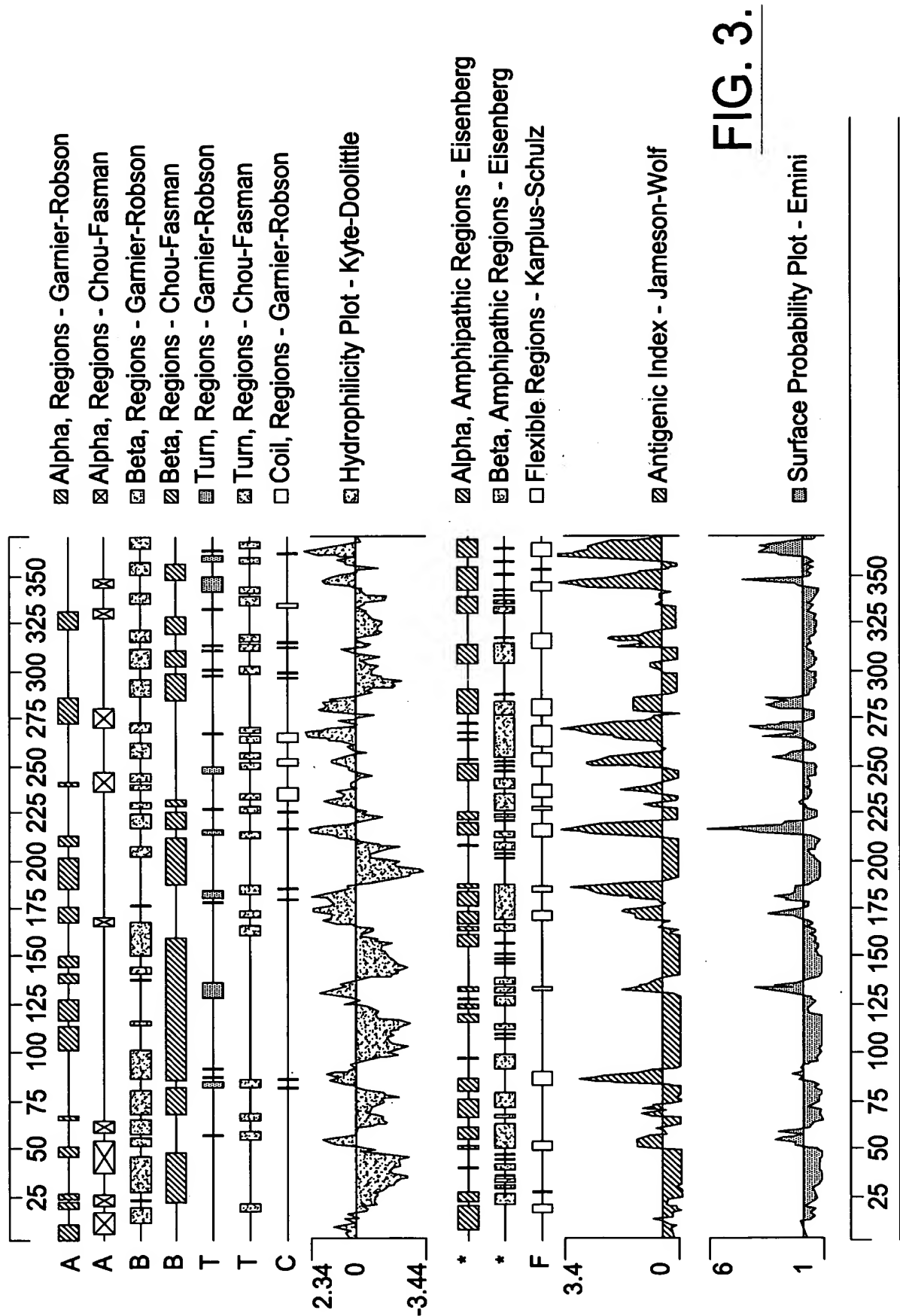
Score: 146.01 Seq: 37 339 Model: 1 269  
 GNILVIWvIcRyRRMRTPMNYFIvNLAvADLLFslftMPFWMvYyvMgg  
 GN+L+ +++++ + +++ ++YF++ L +D+L S ++PF + + ++  
 f1h14926or 37 GNLLISILLVKDKTLHRAPYYFLDLCCSDILRSAICFPFVFNsvKNGS 85  
 RWpFGdfMcrIWmYFDYMNMYASIFFLTcISIDRYLWAICHPMrYmRWMT  
 +W++G++ C+++ ++ +++++ F+L CIS+ RYL AI H Y + T  
 f1h14926or 86 TWYGTLTCKVIAFLGVLSCFHTAFMLFCISVIRYL-AIAHHRFYTKRLT 134  
 pRHRawvMIiIwvMSF1ISMPPFLMFr. WstyrDEneWNmTWcmIyDWP  
 + +++++I+++W++S++++PP+L + +S+ R E++ C++ +  
 f1h14926or 135 FW-TCLAVICMVWTLsvAMAFPPVLDVGTYSFIREEDQ-----CTF-Q-- 175  
 ewMWvWYvILmtiimgFYIPMiIMIFCYwRIYRIaRIWMRMipswQr...  
 +R++ +GF++ + ++L ++Y + ++ + ++  
 f1h14926or 176 ---HRSFR-ANDS-LGMLLLALILLATQLVYLKLIFVHDR---RKMKP 217  
 .....  
 f1h14926or 218 VQFVAASQNWTFHGPgASGQAANWLAGFGRGPTPPTLLGIRQNANTTG 267  
 RRrms....mRrERRivKMliIIMvVFIICWIPYFIvmfMDTLMMwwFCe  
 RRR+ +E+RI++M I+ ++F+ W+PY ++ + +F  
 f1h14926or 268 RRRLLVLDEFKMEKRISRMFYIMTFLFLTLWGPYLVACY-----WRVFAR 312  
 fC. IwrrIWmYIfewLaYvNCpCINPIIY\*  
 ++ + +W++++ INP++  
 f1h14926or 313 GPVVPGGFLT-AAVWMSFAQA-GINPFVC 339

>MILPAT00028/ngf NGF / BDNF / Neurotrophins 3,4, and 6 family of cytokines

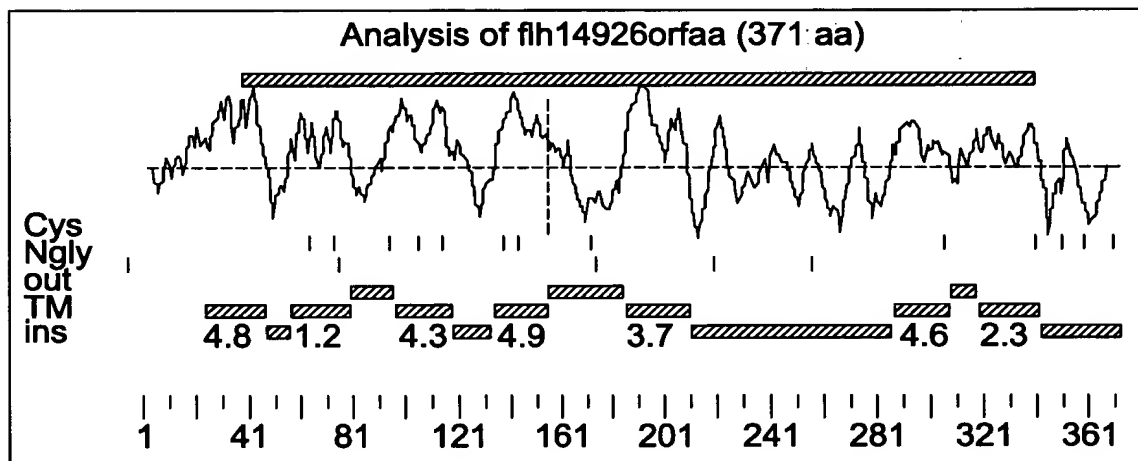
Score: 0.47 Seq: 290 302 Model: 1 13  
 REF xxxxxxxxxxxxxxx  
 \*MSMLFYTMFIsvYF\*  
 M+ LF+T+ +Y+  
 f1h14926or 290 MTFLFLTLWGPYL 302

**FIG. 2.**

4/6



5/6



>flh14926orfaa  
MANYSHAADNILQNLSP LTAFLKLTSLGFIIGVSVVGNLLISILLVKDKTLHRAPYYFL  
DLCCSDILRSAICFPFVNSVKNGSTWTYGTLTCKVIAFLGVLSCFHTAFMLFCISVTRY  
LAIAHHRFYTKRLTFWTCLAVICMVWTLVAMAFPPVL DVGTYSF IREEDQCTFQHR SFR  
ANDSLGFMLLLALILLATQLVYLKLIFFVHDRRKMKPVQFVAAVSQNWTFHGPGASGQAA  
ANWLAGFGRGPTPPTLLGIRQNANTTGRRRLVLDEFKMEKRISRMFYIMTFLFTLWGP  
YL VACYWRVFARGPVVPGGFLTAAVWMSFAQAGINPFVCIFSNRELRRCFSTLLYCRKS  
RLPREPYCVI\*

**FIG. 4.**

6/6

# prositescan will scan one or more sequences against a set of sequence patterns

Database: Release 12.2 of February 1995

Tue Apr 7 18:49:19 1998  
1025 patterns

Query= flh14926orfaa

>PS00001/PDOC00001/ASN\_GLYCOSYLATION N-glycosylation site.  
N[^\*P][ST][^P]

Query:3 NYSH 6

Query:83 NGST 86

Query:182 NDSL 185

Query:227 NWTF 230

Query:264 NTTG 267

>PS00004/PDOC00004/CAMP\_PHOSPHO\_SITE cAMP- and cGMP-dependent protein kinase  
phosphorylation site.  
[RK][2][A-Z][ST]

Query:131 KRLT 134

Query:281 KRIS 284

>PS00005/PDOC00005/PKC\_PHOTPHO\_SITE Protein kinase C phosphorylation site.  
[ST][A-Z][RK]

Query: 80 SVK 82

Query: 93 TCK 95

Query: 130 TKR 132

Query: 178 SFR 180

Query: 266 TGR 268

Query: 342 SNR 344

>PS00006/PDOC00006/CK2\_PHOSPHO\_SITE Casein kinase II phosphorylation site.  
[ST][A-Z][2][DE]

Query: 342 SNRE 345

>PS00008/PDOC00008/MYRISTYL N-myristoylation site.  
G[^\*EDRKHPFYW][A-Z][2][STAGCN][^\*P]

Query: 84 GSTWTY 89

**FIG. 5.**